BILYTICA AI OPERATING SYSTEM

Data Foundation Framework

Five layers that turn scattered systems into data an AI system is allowed to use.

DFF · Foundation

WHAT IT IS FOR

Most AI programmes stall because the data underneath was never contracted, reconciled or certified. DFF makes the foundation an engineered deliverable with a pass mark, not a background clean-up task.

WHO RUNS IT

Run by data engineering and data governance; signed off by the data owner of each domain.

TYPICAL TIMELINE

First certified domain in six to ten weeks; each additional domain runs shorter as the contracts and tests are reused.

STAGES AND EXIT GATES

1. Capture — Inventory every source that feeds the decision, including the spreadsheets nobody lists.

Exit gate: Every source has a named owner, a refresh frequency and a lawful basis for use.

Artefacts: Source register · Lineage sketch · Lawful-basis note

2. Contract — Write a data contract per feed: schema, semantics, freshness, ownership and breakage policy.

Exit gate: Contracts are enforced in the pipeline, not in a document — a breach fails the build.

Artefacts: Data contracts · Contract tests · Breakage runbook

3. Consolidate — Land, conform and reconcile into one governed layer with a single set of keys.

Exit gate: Reconciliation against the system of record is explainable line by line.

Artefacts: Conformed model · Entity resolution rules · Reconciliation report

4. Curate — Define metrics once, classify sensitivity, mask what must be masked, publish for reuse.

Exit gate: One definition per metric, and no unclassified personal data downstream.

Artefacts: Metric dictionary · Classification matrix · Masking policy

5. Certify — Score the domain against the readiness tests and issue an AI-ready certificate with an expiry.

Exit gate: The domain passes freshness, completeness, accuracy and access tests on a schedule, not once.

Artefacts: AI-ready certificate · Quality test suite · Renewal schedule

STANDARDS CROSSWALK

STANDARD	WHAT IT SATISFIES
DAMA-DMBOK	Data quality, metadata, master data and governance knowledge areas.
NDMO	Data management and personal data protection domains, evidenced per domain.
PDPL	Lawful basis, minimisation and masking captured at Capture and Curate.

STANDARD	WHAT IT SATISFIES
NIST AI RMF	Map function — data provenance and validity inputs for AI risk decisions.

WHAT YOU GET

- **Domain data contracts** — Machine-enforced contracts with tests wired into the pipeline.
- **AI-ready certificate** — A dated scorecard per domain with the exact tests that were run.
- **Metric dictionary** — One definition per business metric, with lineage back to source.

WHAT WE REFUSE TO DO

- **Cleaning without a contract** — The same defects return next quarter because nothing upstream changed.
- **A lake with no owners** — Storage grows, trust does not, and every number needs a human to defend it.
- **Certifying once** — A one-off audit ages badly; readiness has to be re-tested on a schedule.

FREQUENTLY ASKED

Do we have to fix all our data first?

No. DFF is run per domain and per use case. You certify the slice a workload depends on, then extend.

What makes a domain AI-ready?

It passes freshness, completeness, accuracy, lineage and access tests repeatedly, and every metric it exposes has one definition.

How does this connect to compliance?

Capture and Curate produce the evidence NDMO and PDPL reviewers ask for, so assurance reuses the same artefacts.

What are the five stages of the Data Foundation Framework?

Capture, Contract, Consolidate, Curate and Certify. Each stage has an exit gate and named artefacts, so a domain only moves forward when the evidence exists.

Book the Data Foundation Assessment

Full framework detail: https://bilytica.com/frameworks/data-foundation-framework/

bilytica.com
info@bilytica.com